DOCUMENT RESUME

ED 323 224 TM 015 427

AUTHOR Nielsen, Lore A.; Turner, Susan D.
TITLE Teaching Writing as a Process: Evaluation for District Decision-Making.
PUB DATE 84
NOTE 18p.
PUB TYPE Journal Articles (080) -- Reports - Research/Technical (143)
JOURNAL CIT Florida Journal of Educational Research; v26 n1 p93-108 Fall 1984

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Attitude Measures; *Decision Making; *Elementary School Students; Elementary School Teachers; Formative Evaluation; Grade 4; Grade 5; Intermediate Grades; Interviews; Pretests Posttests; *Program Evaluation; Program Implementation; Program Improvement; *School Districts; Student Attitudes; Writing (Composition); Writing Evaluation; *Writing Instruction; Writing Teachers
IDENTIFIERS Florida; Writing Samples

ABSTRACT

Formative evaluation of program implementation and student performance was used to assess a new writing program used in 11 elementary schools in Hillsborough County (Florida). Twenty-one teachers and 600 students in grades 4 and 5 participated. Teachers were interviewed, and students completed an attitude survey. In addition, 4 writing samples were evaluated, 2 pre-instruction and 2 post-instruction, from 200 randomly selected students. Fifty students comprised the following subgroups: fourth-grade average, fourth-grade above average, fifth-grade average, and fifth-grade above average. Teachers perceived that the program was implemented as intended, but reported that they needed more specific feedback on their methods. The ranking of student compositions indicated that 76% of the students wrote better after instruction. Overall, students attitudes toward writing were more positive after they participated in the writing instruction program than before they participated. The evaluation also identified specific areas for program improvement. Three data tables are included. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED323224

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Lou M. Carey

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TEACHING WRITING AS A PREOCESS: EVALUATION FOR DISTRICT DECISION-MAKING

Lore A. Nielsen
Susan D. Turner
Hillsborough County Public Schools

TM 015427

BEST COPY AVAILABLE

Florida Journal of Educational Research
Fall 1984, Vol. 26,
No. 1, Pp. 93-108

Teaching Writing as a Process: Evaluation for District Decision-Making

Lore A. Nielsen
Susan D. Turner
Hillsborough County Public Schools

ABSTRACT. New instructional programs should evolve using formative evaluation as a tool for program development. Formative evaluation of program implementation and student performance was used to assess a new writing program used in 11 Hillsborough County elementary schools. Twenty-one teachers and 600 fourth and fifth grade students participated. Teachers were interviewed, 200 randomly selected students' writing samples before and after instruction were evaluated, and students completed an attitude survey. Teachers perceived that the program was implemented as intended but reported that they need more specific feedback on their methods. Students writing skill improved and their attitudes about writing were more positive. Specific areas for program improvement were identified.

Depending on one's perspective, education is currently either basking in the limelight or blenching in the glare of public scrutiny and concern. One area, writing instruction, is receiving attention from a variety of sources. Recently, the Carnegie Foundation for Advancement of Teaching in its report lamented the neglect of writing and concluded, "Teaching students to write clearly and effectively should be a central objective of the school" (Boyer, 1983, p. 91).

Educators in language arts seem prepared to respond to the public concern for renewed emphasis on written expression. Several decades of research, based on theoretical and philosophical considerations, have yielded a variety of generally accepted approaches to

TM015427

teaching writing. The acid test of any of these methods remains, however, whether students improve in their ability to express themselves clearly, correctly, and creatively in their writing.

Along with increasingly sophisticated pedagogy, state-of-the-art technology to assess the effectiveness of writing instruction has also advanced dramatically to provide valid and reliable information (Cooper and Odell, 1977; Diederich, 1974; Myers, 1980; Spandel and Stiggins, 1980). This technology stems from the large body of writing research (Braddock, Lloyd-Jones, and Schoer, 1963), efforts of the Educational Testing Service (Conlon, 1976; Fowles, 1978) and other test publishers (Gardner, Rudman, Karlsen and Merwin, 1983), federal endeavors such as the National Assessment of Educational Progress (1977), and a growing number of state assessments of students' minimum competency in writing (Melton and McCready, 1981).

While the methodology to measure change in writing skills exists and is widely used at the state and national levels, it is not as commonly employed within local districts. Limited financial and personnel resources often make it difficult to conduct writing program evaluations to provide decision-makers with a sound basis to select and modify successful programs.

Despite the difficulties, quantitative measures of writing proficiency need to be used, along with other data sources such as interviews and observations, at the local level as part of both formative and summative evaluations of writing programs. An evaluation paradigm should be developed in which methods are matched to specific questions, selection is made among multiple methods, and emphasis is placed on how information will be used (Patton, 1981; Nielsen and Turner, 1983). This paper presents an example of how one large Florida district evaluated a writing project being piloted in its elementary schools.

Program Description

Under district mandate, several instructional supervisors in Hillsborough County sought to identify and develop an effective model for an elementary writing

program early in the 1982-83 school year. This model was intended to expand the traditional language arts program to include writing in all areas of the curriculum and to be organized using a diagnostic teaching framework. If successful, the program, piloted in eleven schools with 21 teachers and approximately 600 students during the second semester of 1982-83, would eventually be implemented in all 90 district elementary schools.

After reviewing a variety of approaches to teaching writing, the supervisors decided that a systematic, developmental approach would best meet local needs. The program they designed was based on the belief, currently supported by theorists in the field (Emig, 1971; Graves, 1974; Murray, 1968) that composing is a process rather than a product and, therefore, writing instruction should encompass the entire process including prewriting, drafting, editing, revising, and sharing.

Furthermore, since a large number of teachers would be working with students at different grade and ability levels, supervisors decided that the program must be a totally integrated curriculum rather than a collection of writing activities. Therefore, identified writing objectives and instruction were arranged hierarchically from words to sentences, sentences to paragraphs, and paragraphs to longer discourse at four levels of developmental competency. The basis for the pilot program was the manual Teaching Writing: A Developmental Systematic Approach by Evelyn Rothstein and the accompanying Easy Writer activity books (1981).

The theoretical basis for the program, instructional techniques, and management strategies were presented to selected teachers during three days of inservice training spaced between January and April 1983. Teachers were given broad guidelines as well as specific activities that they could tailor to their individual situations. Teachers applied their training in regularly occurring writing lessons in their classes at least three times weekly. Program supervisors frequently visited classrooms to observe and assist teachers.

The purpose of the evaluation of the County Developmental Writing Program was to provide information that could be used to answer the following questions:

1. To what degree was the program implemented in the pilot schools?
2. What modifications would be necessary to improve the program for the subsequent year?
3. As a result of participating in the developmental writing program, did students learn to write better?
4. As a result of participating in the developmental writing program, in what specific ways did student writing change?
5. As a result of participating in the developmental writing program, did student attitudes toward writing change?

Methods

Program evaluation of the elementary writing program consisted of both the implementation process and student product measures. For the implementation evaluation, two evaluators from the Department of Testing and Evaluation conducted 30-45 minute scheduled interviews with all but one program implementor in March 1983. Nineteen classroom teachers, one media specialist, and ten curriculum intervention specialists were interviewed individually. Program-specific interview schedules were developed by the Department of Testing and Evaluation and reviewed by program developers before use. Results were summarized and recommendations were made for modifications in program implementation.

The second portion of the evaluation was designed to investigate each of the student writing quality questions raised for this study. All participating students wrote two compositions within one week in January, 1983. Topics and directions for the compositions were selected and refined by the evaluators and elementary supervisors. Topics were randomly assigned so that half the students wrote on one topic first followed by the second; the other half reversed

this order. Post-program compositions were collected using the same procedures in May 1983.

Sample. Participating students in grades four and five were identified as scoring average, stanines 4, 5, and 6; or above average, stanines 7, 8, and 9, on the spring 1982 CTBS total language test. For the evaluation question about improvement in student writing, a random selection of 200 students, approximately one-third of the population, was made with 50 students in each of the following four grade and ability subgroups: fourth average, fourth above average, fifth average, and fifth above average.

Achievement measures. Four writing samples, two written previous to instruction and two following instruction, were available for each subject. The two pre-samples for each student were paired as were the two post-samples for a total of 400 pairs of compositions. All samples were coded using the students' grade, ability level, student number, and whether the sample was before or after instruction. All other identification was removed from the papers which then were arranged randomly in sets of 15 pairs.

Training readers. Five elementary school teachers who were not part of the pilot program served as readers. During training they were given a review of the program structure and evaluation design. Evaluation criteria were then discussed using material from Diederich (1974) and Myers (1980). The criteria readers were instructed to use in order to judge which of each pair of papers was better were in order of importance: ideas, organization, wording, flavor, usage, spelling, punctuation, and handwriting.

Readers were then given sets of five pairs of student writing for training. All training samples were actual writings done by students in the program who were not part of the 200 chosen subjects. For each pair, readers individually determined which paper was better. After five papers were read, results across readers were tallied and discrepancies discussed. Readers used 20 pairs of papers as part of the initial training; additional training papers were read and discussed following a lengthy break and, as a refresher, at the beginning of the second day of the evaluation. Training continued until there was con-

sistent 80 percent agreement (four out of five readers).

Scoring compositions. Each reader worked with one set of 15 pairs. For each pair, the readers determined which paper was better and recorded their choices on a tally sheet: after all 15 pairs were read, the packet was returned to the evaluator who recorded the choices on a master record sheet and replaced the tally sheet. After a second reader judged the papers without knowing the first reader's choices, the evaluator separated those pairs on which there was agreement: discrepant judgments were returned to the two readers. Differences were reconciled by discussion between readers: if they could not agree, a third reader read the pair in question. The percentage of agreement between readers on the pairs of papers written within a week of each other was 66 percent.

After all pre- and post-pairs had been read separately, the best of each was matched by student identification number resulting in 200 new pairs. Procedures for reading and recording the better of each pair were repeated. The first two readers agreed 71 percent of the time on the pre- and post-samples. All results were subjected to chi-square analyses.

To answer the evaluation questions related to changes in student writing, a sample of 80 pre/post-composition pairs was randomly selected from the 200 subjects used in the general impression ranking. Each pair consisted of the better of the students' pre- and post-writing, with 20 pairs in each of the same four subgroups described previously.

In designing the evaluation, the evaluators and program supervisors examined program materials to determine in what ways student writing might be expected to change as a result of program participation. Six areas were subsequently examined in the feature analysis: fluency (number of words per composition and number of words per t-unit); vocabulary (number of adverbs and adjectives); sentence construction (number of complete sentences, fragments, and run-ons and correct use of initial capital letters and terminal punctuation); sentence variety (number of simple, compound, complex, compound-complex sentences and number

of sentences beginning with verbals or subordinate clauses); paragraph construction (correct use of paragraph format, presence of supports); and editing indications (insertions, substitutions, deletions or text movements).

Three of the elementary teachers who served as readers for the general impression ranking also worked on the feature analysis. Each unit of analysis was defined and exemplified by the evaluator and sample compositions were analyzed. Readers were then assigned categories to analyze and the evaluator checked their work at regular intervals to ascertain its accuracy. All findings were analyzed using either chi-square or t-test statistics, depending on whether the data were nominal or interval.

Measuring attitudes. For the final evaluation question about student attitude towards writing, an attitude survey was developed by the Department of Testing and Evaluation and the elementary supervisors involved in the developmental writing program prior to program implementation. A Likert-type instrument with three categories of response was used. Respondents were asked to indicate their degree of agreement or disagreement with a series of statements about the affective characteristics being examined. The surveys were administered in January before beginning the developmental writing program and again in May after participating in the program for four months. Changes in attitudes were analyzed using t-tests to determine significance.

Results

Program implementation. A summary of the interviews indicated that the Developmental Writing Program was generally implemented according to program guidelines. Most people were satisfied with their participation in the Developmental Writing Program. Generally they were satisfied with most program materials and training sessions and chose to continue participation in the program. As a result of the program, most teachers spent between three and four hours per week in writing instruction, primarily using the Rothstein manual, Easy Writer activities, and various other com-

merical or teacher-made materials. Major program concepts in use included the writing/editing process, Easy Writer activities, and publishing of students' writing. Both district- and school-level support services were provided and were seen by most program teachers as sufficient for their needs.

Program modifications. Some suggestions for program change and some unmet needs were identified. More frequent visitations and more specific feedback on lesson delivery were requested by almost half of the participating teachers.

Improved writing skill. In the general impression ranking of 200 pairs of compositions, 76 percent (152) of the students wrote better post than pre-compositions: that is, about three quarters of the students in the sample improved their writing skill from January to May. Percentages of students who improved for each subgroup are listed in Table 1.

Eighty-six percent of the fifth grade average students improved their compositions followed by fifth grade above average students (82%), fourth grade average students (70%), and fourth grade above average students (66%). A chi-square analysis of the data revealed that all percentages of improvement were significant at the .01 level.

TABLE 1

Percentage of Students Whose Post-instruction Composition was Rated Superior to Their Pre-instruction Composition using a General Impression Criterion

	CTBS Total Language Test Stanines	
Grade Level	Average (4,5,6)	Above Average (7,8,9)
4	70%	66%
5	86%	82%

Specific writing changes. Results of the subsequent analyses of specific characteristics of student writing are summarized in Table 2. Fluency factors

TABLE 2

A Comparison of Writing Skill on Selected Factors from Pre-Instruction to Post-Instruction Compositions

Factors	Writing Sample	Grade Level 4th	5th	Total Group
Fluency				
Number of Words per Composition	Pre	116	133	124
	Post	134*	128	131
Number of Words per t-Unit	Pre	8.7	9.5	9.1
	Post	9.3	9.1	9.2
Vocabulary				
Number of Adverbs	Pre	5.5	5.9	5.7
	Post	7.0*	5.3	6.2
Number of Adjectives	Pre	6.4	6.9	6.7
	Post	8.1	8.9*	8.5*
Sentence Construction				
Number of Complete Sentences	Pre	8.6	10.1	9.4
	Post	10.7*	10.1	10.4
Number of Run-Ons	Pre	1.1	0.9	1.0
	Post	1.1	1.1	1.1
Number of Fragments	Pre	0.6	0.5	0.5
	Post	0.3	0.3	0.3
Percentage Who Correctly Use Capitals	Pre	65	90	78
	Post	73	88	80*
Percentage Who Correctly Use Terminal Punctuation	Pre	83	90	86
	Post	83	88	85

* $p < .01$

TABLE 2 Continued

Factors	Writing Sample	Grade Level 4th	5th	Total Group
Sentence Variety				
Number of Simple Sentences	Pre	5.9	6.2	6.1
	Post	7.3	6.2	6.8
Number of Compound Sentences	Pre	0.9	1.1	1.0
	Post	1.2	1.0	1.1
Number of Complex Sentences	Pre	1.7	2.6	2.1
	Post	1.7	2.7	2.2
Number of Compound-Complex Sentences	Pre	0.2	0.7	0.3
	Post	0.4	0.6	0.5
Number of Variant Beginnings	Pre	0.7	1.2	0.8
	Post	0.8	1.4	0.9
Paragraph Construction				
Percentage Who Correctly Use Paragraph Format	Pre	48	75	61
	Post	73	82	78
Percentage Who Use Topic Sentences	Pre	50	68	59
	Post	55	70	63
Number of Support Sentences Used	Pre	5.4	4.2	4.8
	Post	5.3	5.7*	5.5
Percentage Who Use Concluding Sentences	Pre	40	75	58
	Post	75	90	83

* $p < .01$

TABLE 2 Continued

Factors	Writing Sample	Grade Level 4th	5th	Total Group
Editing				
Percentage Making Insertions	Pre	3	10	6
	Post	25	22	24
Percentage Making Substitutions	Pre	10	26	38
	Post	5	15	21
Percentage Making Deletions	Pre	5	15	10
	Post	18	18	18
Percentage Making Movements	Pre	0	2	3
	Post	0	0	0

identified include words per composition and words per t-unit. Fourth grade students significantly increased the number of words per composition from an average of 116 words on the pre-sample to an average of 134 words on the post-sample. For the total group, the trend on both words per composition and words per t-unit was positive although neither area reached statistical significance.

Vocabulary measures included the number of adjectives and number of adverbs per composition. Fourth grade students significantly increased the average number of adverbs used from 5.5 on the pre-sample to 7.0 on the post-sample. The increase in the average number of adjectives used was significant for fifth grade students and for the total group, 6.9 to 8.9 adjectives and 6.7 to 8.5 adjectives, respectively. Again, changes were in a positive direction.

Five measures of sentence construction characteristics were identified. First, the average number of complete sentences, run-ons, and sentence fragments were tallied. Then, the percent of students who correctly used capitals and terminal punctuation were calculated. Fourth grade students significantly increased the number of complete sentences from the

pre- to post-composition from 8.6 to 10.7. The increase in the percentage of fourth grade students and the total group that correctly used initial capitalization was also significant, 65 percent to 73 percent and 78 percent to 80 percent, respectively.

Sentence variety as indicated by types of sentences--simple, compound, complex and compound-complex -- and variant sentence beginnings were summarized. Although changes observed were in a positive direction, none reached statistical significance.

Factors related to paragraph construction were percentages of pupils using correct paragraph format, having a topic sentence, and having a concluding sentence, as well as the number of topic supports included in the paragraph. The percentage of students including a topic sentence increased significantly from 68 percent to 70 percent for fifth grade students. Fifth grade students also significantly increased the number of supports they used in their pre- to post-compositions from 4.2 to 5.7. Other changes were in a positive direction although none were significant.

Finally, editing indications were summarized. There were no significant changes in percentages of students making insertions, substitutions, deletions, or movements, and the percentages of students using these editing techniques were generally small (ranging from none to 25%).

Student attitudes. The final area of investigation was in the affective domain. Overall, student attitude toward writing was more positive after participating in the Developmental Writing Program. Pre-post analysis of student responses to the attitude survey indicated a significant increase in positive attitudes toward writing. T-test statistics for the 481 subjects are listed in Table 3. Of a possible 60-point total (i.e., if responses to all 20 items were positive), participating students averaged 45.8 on the pre-survey and 47.1 on the post-survey. Although this difference (1.3) was slight, it was consistent among students in both degree and direction.

TABLE 3

Student Attitudes Compared Before and After Participating in the Writing Program

Attitude Survey	$\bar{X}$	SD	T-Value
Pre-Instruction	45.8		
Post-Instruction	47.1	5.99	4.75*

$p < .000$
$N = 481$ students

Discussion

As a result of the project evaluation of the Developmental Writing Program in selected Hillsborough County elementary schools, decision-makers learned the following: the project was generally implemented according to guidelines, and teachers were satisfied with their participation: students' writing production was significantly better (76% improved from pre-instruction to postinstruction compositions) and their attitudes towards writing were more positive after program participation. Students improved in some of the specific factors examined in the feature analysis, although most changes did not reach statistical significance. Based on these findings, the writing program was continued for 1983-84. The program has been expanded to involve additional personnel at the pilot schools and has been introduced to more than 20 new schools.

Concomitant to program continuation, the evaluation design changed to meet different evaluation concerns for 1983-84. With initial confirmation of the program's feasibility, a pretest-posttest control group design has been deemed appropriate to measure whether changes in students' writing are due to program intervention or to other factors such as maturation or by-school variations. Also, the type of scoring will change from use of the gross ranking "better" to a holistic score one to six, which will be

summed for the two pre- and two post-writing samples. This procedure will allow application of the t-test to results and increase interrater reliability. CTBS reading and language scores will be examined to determine how the writing program affects more traditional language arts skills. The feature analysis will not be repeated. Moreover, instrumentation, especially measurement of student and teacher attitudes, has been refined through standard test development procedures. Process data collection has become more specific through use of the Concerns-Based Adoption Model (Loucks, Newlove and Hall, 1975) which will be used to determine to what extent the program is used by teachers. Finally, systematic observation of classroom teaching using locally developed instruments will confirm and expand interview findings.

In addition to the effect of the program evaluation on program continuation and subsequent evaluation design, several less direct implications emerge from this study. First, the evaluation serves as an example of how staff development efforts can be evaluated by observing changes in teacher behavior and by measuring student outcomes. This issue is especially timely in Florida, where the legislature has mandated extensive product evaluation of district staff development offerings.

A second implication can be drawn from the fact that classroom teachers were actively involved in analyzing student writing. By training and using teachers as readers, their professional understanding and skills were enhanced, especially their ability to evaluate student writing. Furthermore, these teachers returned to their schools with their new expertise and helped disseminate program goals more widely across the district.

In conclusion, the evaluation of the Developmental Writing Program in Hillsborough County is an example of how, with careful planning and marshalling of resources, sound program evaluation of writing can take place at the district level to assist in local decision-making. By incorporating evaluation procedures into new programs from the planning stages onward, such programs gain in credibility and are more likely to receive continued district support.

Finally, evaluation design can evolve with each year of program development, building on prior evaluation foundations and responding to changing program needs.

References

Boyer, E. (1983). High school: A report on secondary education in America. New York: Harper and Row.

Braddock, R., Lloyd-Jones, R., and Schoer, L. (1963). Research in written composition. Urbana, IL: National Council of Teachers of English.

Conlon, G. (1976). How the essay in the CEEB English test is scored. Princeton, NJ: Educational Testing Service.

Cooper, C., and Odell, L. (1977). Evaluating writing: Describing, measuring, judging. Urbana, IL: National Council of Teachers of English.

Diederich, P. (1974). Measuring growth in English. Urbana, IL: National Council of Teachers of English.

Emig, J. (1971). The composing processes of twelfth graders. (Research Report No. 13). Urbana, IL: National Council of Teachers of English, Committee on Research, ED 038 252.

Fowles, M. E. (1978). Manual for scoring the writing sample: Analytical scoring, holistic scoring. Princeton, NJ: Educational Testing Service.

Gardner, E., Rudman, H., Karlsen, B., Merwin, J. (1983). Stanford Writing Assessment Program. NY: Harcourt, Brace, Jovanovich.

Graves, D. H. (1974). Children's writing: Research directions and hypotheses based upon an examination of the writing process of seven-year-old children. (Doctoral dissertation, State University of New York at Buffalo, 1973). Dissertation Abstracts International. 34, 6255A. (University Microfilms No. 74-08375) ED 095 586.

Loucks, S., Newlove, B., and Hall, G. (1975). Measuring levels of use of the innovation: A manual for trainers, interviewers, and raters. Austin, TX: University of Texas.

Melton, V., and McCready, M. (1981). Survey of large-scale writing assessment programs. Ruston, LA: Louisiana Technological University.

Murray, D. (1968). A teacher teaches writing. Boston: Houghton-Mifflin Co.

Myers, M. (1980). A procedure for writing assessment and holistic scoring. Urbana, IL: National Council of Teachers of English.

National Assessment of Educational Progress. (1977). Write/rewrite: An assessment of revision skills. Denver, CO: NAEP.

Nielsen, L. and Turner, S. (1983). Program evaluation as an evolutionary process. Evaluation Review, 1, 399-405.

Patton, M. Q. (1981). Creative evaluation. Beverly Hills: Sage Publications.

Rothstein, E. (1981). Teaching writing: A developmental systematic approach. Nyack, NY: The Write Track.

Spandel, V., & Stiggins, R. (1980). Direct measures of writing skill: Issues and applications. Portland, OR: Northwest Regional Educational Laboratory.

AUTHORS

LORE A. NIELSEN, Supervisor of Staff Development, Hillsborough County School District, 901 E. Kennedy Blvd., P. O. Box 3408, Tampa, Florida 33601

SUSAN D. TURNER, Supervisor of Evaluation, Hillsborough County School District, 901 E. Kennedy Blvd., P. O. Box 3408, Tampa, Florida 33601